

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/CN04/001435

International filing date: 10 December 2004 (10.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: CN
Number: 200310117354.6
Filing date: 11 December 2003 (11.12.2003)

Date of receipt at the International Bureau: 09 March 2005 (09.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

证 明

本证明之附件是向本局提交的下列专利申请副本

申 请 日: 2003. 12. 11

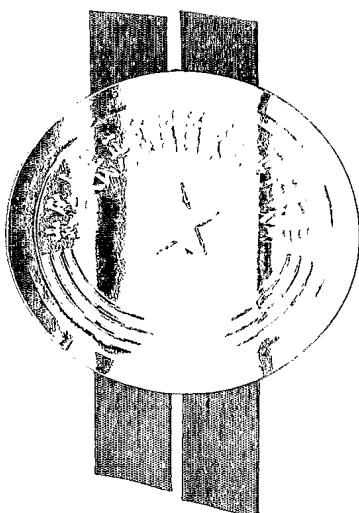
申 请 号: 2003101173546

申 请 类 别: 发明

发明创造名称: 一个蛋白质家族的抗菌用途

申 请 人: 中国人民解放军军事医学科学院放射医学研究所

发明人或设计人: 赵士富、赵彦林、张咏、唐芊芊



中华人民共和国
国家知识产权局局长

王景川

2005 年 1 月 11 日

权 利 要 求 书

- 1、一个蛋白质家族在医药领域、化妆品、食品、畜禽水产饲料、种子包衣以及提高转基因动植物抗病虫害的能力中的应用，所说的蛋白质家族包括具有下列 GenBankTM 注册号的蛋白质：CAC11117, AAK18748, AAH59661, XP_320383, NP_648757, NP_496804, B88349, CAB89006, AAB68201, NP_566324, NP_705421, EAA17587, EAA47634, EAA34843。
- 2、根据权利要求 1 所述的用途，其特征在于用于制备细菌性感染疾病的治疗药物。
- 3、根据权利要求 1 所述的用途，其特征在于用于筛选权利要求 1 所述蛋白质家族中蛋白质的配体。
- 4、根据权利要求 1 所述的用途，其特征在于用于化妆品添加剂。
- 5、根据权利要求 1 所述的用途，其特征在于用于食品防腐剂。
- 6、根据权利要求 1 所述的用途，其特征在于用于畜禽或水产饲料添加剂。
- 7、根据权利要求 1 所述的用途，其特征在于用于种子包衣。
- 8、根据权利要求 1 所述的用途，其特征在于用于提高转基因植物或动物抗病虫害的能力。
- 9、根据权利要求 1 所述的用途，其特征在于用于鲜花保鲜。

说 明 书

一个蛋白质家族的抗菌用途

技术领域

本发明涉及一个蛋白质家族的用途，具体地说涉及一个蛋白质家族的抗菌用途。

背景技术

20多年来，医药界没有发现任何新的抗菌素家族，但是抗药性现象的发展却极为迅速，人们甚至可以用肉眼观察到细菌的基因突变。更为可怕的是，对付金葡萄球菌的障碍刚刚攻克，美国医生就发现了对可有效抑制细菌的最新抗菌素—万古霉素具有抗性的菌株。致病菌的抗药性问题已经日益严重地威胁着人们的健康。寻找全新类型的抗生素是解决抗药性问题的一条有效途径。抗菌肽作为生物天然免疫的活性分子，被各种生物用来抵御来自外界病菌的侵染，在生物界中广泛存在。人为分离出来的天然抗菌肽表现出抗菌活性高，抗菌谱广，种类多，可供选择的范围广，靶菌株不易产生抗性突变等特点，因而在医药工业、食品工业和农业上被认为有着广阔的应用前景。抗菌肽是指由基因编码，在核糖体上合成的相对分子质量通常在10kDa以下，具有抗菌活性的多肽类物质，也叫多肽抗生素。大部分抗菌肽具有热稳定性，在100℃下加热10~15min仍能保持其活性。抗菌肽对较大的离子强度和较高或较低的pH值均具有较强的抗性。多数抗菌肽的等电点大于7，表现出较强的阳离子特征。同时，部分抗菌肽尚具备抵抗胰蛋白酶或胃蛋白酶水解的能力。此外，研究发现不同家族的抗菌肽之间在序列上极少存在同源性，而同一家族的不同成员之间序列上却存在高度的保守性，这意味着其功能也非常保守。抗菌肽除了具有抗细菌或真菌的作用，有些还具有抗原虫、病毒或癌细胞的功能。抗菌肽在生物界中广泛存在。迄今为止，人们已经从细菌、真菌，到两栖类、昆虫、高等植物、哺乳动物、直至人类体内发现了多达700种以上的多肽抗生素。目前，已有多种抗菌肽正在进行临床前的可

行性研究；此外，抗菌肽转基因动物、转基因植物，抗菌肽在食品防腐，鲜花保鲜，化妆品，种子包衣和动物饲料添加剂等方面的应用研究也正在进行之中。

鉴于抗菌肽在医药工业、食品工业和农业上有着广阔的应用前景，通过各种途径从生物界中寻找新型的天然抗菌肽就成为目前世界范围内热点之一。

本发明的发明人通过试验首次发现了一种蛋白质，将其命名为 Glyrichin 蛋白，然后利用生物信息学技术，通过同源性分析，发现了一个包含多个蛋白质在内的蛋白质家族，Glyrichin 蛋白家族。目前，这些蛋白质的功能还不清楚，也没有这些蛋白质的抗菌特性的报道。

发明内容

本发明公开了一个天然蛋白质家族的用途。这个蛋白质家族中的蛋白质包括来自酵母菌、线虫、果蝇、按蚊、斑马鱼、小鼠到人等物种的高度同源体蛋白质，具有如表 1 中序列 1~14 所示的氨基酸序列。

本发明所述的蛋白质家族的用途包括以下方面：

- (1) 直接做为人或畜细菌性感染疾病的治疗药物；
- (2) 用于筛选促进或拮抗 Glyrichin 蛋白功能的抗体、多肽或其它配体；
- (3) 用于化妆品添加剂；
- (4) 用于食品防腐剂；
- (5) 用于畜禽或水产饲料添加剂；
- (6) 用于种子包衣；
- (7) 用于转基因植物或动物以提高其抗病虫害的能力。
- (8) 用于鲜花保鲜。

本发明所述的蛋白质家族包括下列蛋白质：

表 1 Glyrichin 蛋白家族包含的蛋白质

序号	物种	来源	GenBank™ 注册号
1	人	Homo sapiens	CAC11117

2	小鼠	<i>Mus musculus</i>	AAK18748
3	斑马鱼	<i>Danio rerio</i>	AAH59661
4	按蚊	<i>Anopheles gambiae</i>	XP_320383
5	果蝇	<i>Drosophila melanogaster</i>	NP_648757
6	线虫	<i>Caenorhabditis elegans</i>	NP_496804
7	线虫	<i>Caenorhabditis elegans</i>	B88349
8	芽殖酵母	<i>Schizosaccharomyces pombe</i>	CAB89006
9	酿酒酵母	<i>Saccharomyces cerevisiae</i>	AAB68201
10	拟南芥	<i>Arabidopsis thaliana</i>	NP_566324
11	疟原虫	<i>Plasmodium falciparum</i> 3D7	NP_705421
12	疟原虫	<i>Plasmodium yoelii</i> <i>yoelii</i>	EAA17587
13	稻瘟病菌	<i>Magnaporthe grisea</i>	EAA47634
14	脉孢菌	<i>Neurospora crassa</i>	EAA34843

本发明所述的蛋白质家族还包括具有在上述蛋白质的氨基酸序列基础上发生氨基酸残基的插入、取代、缺失、添加等所产生的，具有与原序列相同功能的序列的蛋白质。

发明人利用抑制性差减杂交的方法，对 LTC (long term culture) 培养前后的小鼠骨髓基质细胞差异表达基因进行了大量筛选，获得了 131 个差异表达的 EST 克隆。生物信息学分析证实，这些克隆代表了 26 种已知或部分已知功能的基因和 7 条全新的基因，其中 5 条具有完整的开放阅读框架，3 条已在 GenBank 注册，鼠源的 mGlyrichin 就是其中之一，其基因全长 cDNA 序列 521bp，其 GenBank™ 注册号为 AY028425，参见《实验血液学杂志》2002 第 10 期第 177-182 页。

在得到鼠源 mGlyrichin 后，发明人根据鼠源 Glyrichin 基因设计一对简并引物，通过 PCR 扩增人胎肝 cDNA 文库，获得人源 Glyrichin 完整的 ORF 序列共 240bp，以及根据 ORF 序列推测的氨基酸序列。发明人利用生物信息学手段对序列进行了详细分析，进行了 Glyrichin 家族成员间的同源性比较—多序列对齐。Glyrichin 家族各成员的序列均来自 GenBank，它们的注册号分别为：人 *Homo sapiens*-CAC11117、小鼠 *Mus musculus*-AAK18748、斑马鱼 *Danio rerio*-AAH59661、按蚊 *Anopheles gambiae*-XP_320383、果蝇

Drosophila melanogaster-NP-648757、线虫 Caenorhabditis elegans-B88349 和 NP-496804、芽殖酵母 Schizosaccharomyces pombe-CAB89006、酿酒酵母 Saccharomyces cerevisiae-AAB68201、拟南芥 Arabidopsis thaliana-NP-566324 和疟原虫 Plasmodium falciparum 3D7-NP-705421 和 Plasmodium yoelii yoelii-EAA17587 等。用 DNASTAR Inc. 公司的分析软件 MegAlign 作图，多序列比对图如附图 1 所示，反映出 Glyrichin 家族各成员的共同特点是存在一个高度保守的、大约 60aa 长度的核心结构域，在该结构域的上游或下游的序列则同源性较低。到目前为止，该家族所有成员功能均未知。

为研究 Glyrichin 家族各成员的亲缘关系，对这 12 条序列进行了进化树分析，如附图 2 所示，显示出 Glyrichin 基因从低等到高等的进化历程及各成员间亲缘关系的远近。Glyrichin 家族成员间保守的甘氨酸富含区 (Glyrich domain) 如附图 3 所示。Glyrich 结构域在 Glyrichin 家族不同成员间的序列一致性如附图 4 所示，可以看出每两个成员间的亲缘关系的远近。核酸序列比较结果显示 hGlyrichin 基因定位于人染色体 20q11.21 区域，由两个外显子组成，编码 79 个氨基酸的小分子蛋白。氨基酸水平 BlastP 同源性比较显示该基因在进化上十分保守，存在于从真菌（如脉孢菌、酵母菌等）、植物（如拟南芥等）、疟原虫、线虫、果蝇、按蚊、斑马鱼、小鼠到人等物种中，组成 Glyrichin 同源性蛋白家族；在不同种属间显示出很高的同源性，已知人源 Glyrichin 与鼠源相比 100% 同源，与斑马鱼 Glyrichin 相比具有 90% 同源性，与果蝇的同源性有 62%，与酵母的同源性也有 46%（图 1）；家族内不同成员间的亲缘关系以进化树表示（图 2）。hGlyrichin 的氨基酸序列特征见表 2，为 Protean 分析结果。

表 2 hGlyrichin 的序列特征

项目	内容
分子量	8181.9
多肽长度 (aa)	79
等电点	9.36
电荷数 (pH7.0)	4.79
Gly 含量 (%)	21.52

Glyrichin 家族成员的序列特征见表 3。发明人进行了 hGlyrichin 二级结构预测, hGlyrichin 的结构特点见附图 5, 为 SignalP 分析结果。

表 3 Glyrichin 家族成员的序列特征

物种属名	多肽长度(aa)	等电点	电荷数(pH7.0)	甘氨酸含量(%)
Homo sapiens	79	9.36	4.79	21.52
Mus musculus	79	9.36	4.79	21.52
Danio rerio	80	9.81	4.82	21.2
Anopheles gambiae	128	9.73	10.08	14.1
Drosophila melanogaster	79	9.66	4.82	20.3
Caenorhabditis elegans	145	10.79	12.01	14.48
Schizosaccharomyces pombe	107	10.04	4.05	13.08
Saccharomyces cerevisiae	112	9.74	5.04	13.27
Plasmodium falciparum 3D7	168	9.66	5.15	15.94
Plasmodium Yoelii Yoelii	167	9.16	4.15	15.28
Magnaporthe grisea	107	11.75	8.24	13.08
Arabidopsis thaliana	74	9.84	6.18	14.86

鉴于 Glyrichin 家族各成员进化上高度保守的特点以及其功能未知的事实, 发明人对该基因家族的功能开展研究, 拟证实 Glyrichin 家族成员的功能及其用途。

目前已知 Glyrichin 家族共包括 14 个成员。这 14 个成员分布在从真菌到人的所有基因组已完成测序的物种中。本领域的普通技术人员可以轻易地据此推断出它也存在于其他基因组尚未完成测序的物种中。Glyrichin 家族各个成员结构上共同的特点是存在一个甘氨酸 (Glycine) 富含区, 长度在 59-68 个氨基酸之间。该区域基本涵盖了人源和鼠源 Glyrichin 全部的序列。存在于其他物种中的 Glyrichin 除了共同的甘氨酸富含区外, 还在此保守区的氨基端或羧基端存在其他序列 (如信号肽等), 但不具有保守性 (见图 3、图 4)。这说明甘氨酸保守区决定了 Glyrichin 家族成员最基本的生物活性; 也从另一方面说明生物在进化过程中逐渐淘汰了冗余序列, 使分子变得更短小、结构更紧凑。

在获得人源 Glyrichin 多核苷酸序列后, 发明人以人源和鼠源 Glyrichin 为例, 对该基因家族的功能开展研究, 结果证明人源和鼠源

Glyrichin 是一个在人和小鼠多种组织中广泛表达的具抗菌活性的天然免疫分子，是一种天然存在的抗菌肽，有望发展成为在医药工业、食品工业、化妆品行业、农业、畜牧业和渔业等领域具有广泛用途的新型多肽抗生素。这也是 Glyrichin 家族各个成员的基本功能和活性。

表 2 的统计数据表明人/鼠源 Glyrichin 多肽的分子量小于 10kDa，等电点为 9.36，在 pH 值为 7 的条件下带 4.79 个正电荷。表 3 的 SignalP 分析表明 h/mGlyrichin 分子中存在一个强疏水区。Glyrichin 家族其他成员的等电点均大于 7，在 pH=7 时都带正电荷，且都存在强的疏水区域（图 5）。上述结果表明 Glyrichin 家族具有已知抗菌肽典型的结构特征。对人源 Glyrichin 进行 Northern blot 分析和体外转录与翻译分析，以 $[\alpha - ^{32}\text{P}]$ dATP 标记的 hGlyrichin 为探针检测其转录本大小和表达谱，结果见图 6A。所用的 4 种肿瘤细胞株分别为 HepG2、HeLa、HEK293、Jurkert，显示人 Glyrichin 基因在所测试的 4 种不同组织来源的肿瘤细胞株中都有表达，提示它可能是一种广泛表达的天然抗菌肽。同时揭示只存在一个转录本，且大小约为 600bp。将 hGlyrichin 基因重组到含有 T7 启动子的 pT7 载体中，进行人源 Glyrichin 体外转录和翻译，hGlyrichin 蛋白体外翻译后的大小约为 8.8kDa，同理论推断的结果一致，表明该基因可以在体外正常地转录和翻译，见图 6B。

接下来发明人使用两种方法测试抗菌活性：将 hGlyrichin 基因导入细菌中然后诱导表达观察抑菌效应和分离纯化 Glyrichin 蛋白求最小抑菌浓度。首先，将 hGlyrichin 不带任何标签（即不同任何蛋白融合）重组到原核表达载体 pET-22b (+) 中，将 pET-22b (+)-Glyrichin 转化到大肠杆菌 BL21 中，比较 IPTG 存在和不存在两种情况下细菌生长情况。Glyrichin 在不加 IPTG 诱导剂时不表达，而只在培养液中加入 IPTG 诱导后才大量表达；进一步地，在诱导条件下，如果 Glyrichin 不具有抗菌活性，则同样不会影响细菌生长；但是如果 Glyrichin 是抗菌肽，则会明显影响细菌生长；在后一种情况下，随着培养时间延长，IPTG 不断消耗，Glyrichin 的表达逐渐停止，理论上此时细菌的生长应逐渐恢复。每隔 45 分钟测量 OD600 一次，共测量

Glyrichin 是一个在人和小鼠多种组织中广泛表达的具抗菌活性的天然免疫分子，是一种天然存在的抗菌肽，有望发展成为在医药工业、食品工业、化妆品行业、农业、畜牧业和渔业等领域具有广泛用途的新型多肽抗生素。这也是 Glyrichin 家族各个成员的基本功能和活性。

表 2 的统计数据表明人/鼠源 Glyrichin 多肽的分子量小于 10kDa，等电点为 9.36，在 pH 值为 7 的条件下带 4.79 个正电荷。表 3 的 SignalP 分析表明 h/mGlyrichin 分子中存在一个强疏水区。Glyrichin 家族其他成员的等电点均大于 7，在 pH=7 时都带正电荷，且都存在强的疏水区域（图 5）。上述结果表明 Glyrichin 家族具有已知抗菌肽典型的结构特征。对人源 Glyrichin 进行 Northern blot 分析和体外转录与翻译分析，以 $[\alpha - ^{32}\text{P}] \text{dATP}$ 标记的 hGlyrichin 为探针检测其转录本大小和表达谱，结果见图 6A。所用的 4 种肿瘤细胞株分别为 HepG2、HeLa、HEK293、Jurket，显示人 Glyrichin 基因在所测试的 4 种不同组织来源的肿瘤细胞株中都有表达，提示它可能是一种广泛表达的天然抗菌肽。同时揭示只存在一个转录本，且大小约为 600bp。将 hGlyrichin 基因重组到含有 T7 启动子的 pT7 载体中，进行人源 Glyrichin 体外转录和翻译，hGlyrichin 蛋白体外翻译后的大小约为 8.8kDa，同理论推断的结果一致，表明该基因可以在体外正常地转录和翻译，见图 6B。

接下来发明人使用两种方法测试抗菌活性：将 hGlyrichin 基因导入细菌中然后诱导表达观察抑菌效应和分离纯化 Glyrichin 蛋白求最小抑菌浓度。首先，将 hGlyrichin 不带任何标签（即不同任何蛋白融合）重组到原核表达载体 pET-22b (+) 中，将 pET-22b (+)-Glyrichin 转化到大肠杆菌 BL21 中，比较 IPTG 存在和不存在两种情况下细菌生长情况。Glyrichin 在不加 IPTG 诱导剂时不表达，而只在培养液中加入 IPTG 诱导后才大量表达；进一步地，在诱导条件下，如果 Glyrichin 不具有抗菌活性，则同样不会影响细菌生长；但是如果 Glyrichin 是抗菌肽，则会明显影响细菌生长；在后一种情况下，随着培养时间延长，IPTG 不断消耗，Glyrichin 的表达逐渐停止，理论上此时细菌的生长应逐渐恢复。每隔 45 分钟测量 OD600 一次，共测量

20 次。以时间为横轴, OD600 对数值为纵轴做生长曲线。转化了 pET-22b (+) 空载体的 BL21 在 0.5mM IPTG 诱导和不诱导条件下绘制半对数生长曲线见图 7A, 转化了不具抗菌活性的无关基因 pET-22b (+)-UBF 和 pET-22b (+)-PTP 的 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线见图 7B, 转化了待测的 pET-22b (+)-hGlyrichin 阳性克隆 1 和 8 的 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线见图 7C, 待测的 pET-22b (+)-hGlyrichin 阳性克隆 1 的两次重复试验的曲线图见图 7D。

单转空载体 pET-22b 的菌株在加入和不加入 IPTG 两种情况下细菌的生长曲线基本吻合 (图 7A); 在单转两个已知不具有抗菌活性基因 UBF 和 PTP 的情况下, 加入和不加入 IPTG 两种情况下细菌的生长同样不受影响 (图 7B); 但是, 在转入 Glyrichin 后, 加入和不加入 IPTG 两种情况下细菌的生长曲线明显不同, 在前一种情况下细菌的生长明显受到抑制 (图 7C, D); 从图中还可以看出, 随着培养时间延长 IPTG 消耗殆尽, 细菌的生长又得以恢复, 而其他测试基因都没有此现象。上述结果充分说明这种抑菌效应是 Glyrichin 蛋白本身的作用。在接下来的另一组实验中, hGlyrichin 基因被重组到 pGEX-4T2 原核表达载体中, 然后转化到 BL21 中诱导表达 GST-hGlyrichin 融合蛋白 (图 8); 首先根据小鼠 Glyrichin 的 cDNA 序列设计引物, 以人胎肝 mRNA 进行 RT-PCR 扩增, 纯化后 PCR 产物经测序和限制性内切酶消化后在 T4 连接酶的作用下, 插入到用同样内切酶消化后的 pGEX4-4T2 原核表达载体中, 经转化 JM109 大肠杆菌后, 以 IPTG 进行体外诱导表达 5 小时, 用全菌体蛋白进行 SDS-PAGE 电泳。然后, 纯化游离 hGlyrichin 表达产物并检测其杀菌活性, 将上所述诱导表达得到 GST-Glyrichin 融合蛋白, 经 GST 纯化柱纯化后用 Enterokinase 切割, 得到游离的 hGlyrichin 蛋白, 利用 96 孔板法检测多肽的抗菌活性, 测定最小抑菌浓度 (MIC), 结果表明 GST-Glyrichin 融合蛋白可以在大肠杆菌菌株 JM109 中大量表达, 不能抑制细菌生长, 不具有抗菌活性, 游离的 Glyrichin 蛋白具有在较低浓度下抑制革兰氏阳性菌和革兰氏阴性菌的生长, 具有抗菌活性, 见表 4。

表 4 游离 hGlyrichin 蛋白的最小抑菌浓度 (MIC)

	菌株名称	最小抑菌浓度(ug/ml)
革兰氏阴性菌	大肠杆菌 DH5 α	0.2
革兰氏阳性菌	枯草杆菌 DB430	3

在本发明中，术语“Glyrichin 家族”是指具有 Glyrich 结构域且在该结构域区段氨基酸同源性相互间在 30% 以上的多肽，包括本发明中已经提到的 14 个多肽和其他尚未报道的多肽。以进化生物学的观点，它们来自于同一个祖先。术语“Glyrich 结构域”特指表 1 中序列 1~14 中最为保守的一段多肽，其氨基酸长度在 59~68 之间，等电点大于 7，pH7.0 条件下带正电荷，Glycine 含量在 20 种氨基酸中含量为最高，并且含有至少一个疏水区的一段多肽。它们的序列及结构特点如图 4 和图 5，表 2 和表 3 所界定。

术语“Glyrichin 多肽”指具有 Glyrichin 蛋白抗菌活性的表 1 中序列 1~14 的多肽。该术语还包括具有与 Glyrichin 蛋白相同功能的表 1 中序列 1~14 的变异形式。这些变异形式包括（但并不限于）：若干个（通常为 1~20 个，较佳地为 1~10 个，更佳地 1~5 个，最佳地 1~3 个）氨基酸的缺失、插入和/或取代，以及在 C 末端和/或 N 末端添加一个或数个（通常为 20 个以内，较佳地为 10 个以内，更佳地为 5 个以内）氨基酸。例如，在本领域中，用性能相近或相似的氨基酸进行取代时，通常不会改变蛋白质的功能。又比如，在 C 末端和/或 N 末端添加一个或数个氨基酸通常也不会改变蛋白质的功能。该术语还包括 Glyrichin 蛋白的活性片段和活性衍生物。

该多肽的变异形式包括：同源序列、保守性变异体、等位变异体、天然突变体、诱变突变体、在高或底的严紧度条件下能与 Glyrichin 多肽的抗血清获得的多肽和蛋白。本发明还提供了包含 Glyrichin 多肽或其片段的融合蛋白。除了几乎全长的多肽外，本发明还包括了 Glyrichin 多肽的可溶性片段。通常，该片段具有 Glyrichin 多肽序列的至少约 25 个连续氨基酸，通常至少约 35 个连续氨基酸，更佳地至少约 45 个连续氨基酸，最佳地至少约 60 个连续氨基酸。应理解，本发明的多肽并不限于上述例举的代表性的多肽。

修饰（通常不改变一级结构）形式包括：体内或体外的多肽的化学形式

如乙酰化或羧基化。修饰还包括糖基化，如那些在多肽的合成和加工或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖基化的酶（如哺乳动物的糖基化酶或去糖基化酶）而完成。修饰形式还包括具有磷酸化氨基酸残基（如磷酸丝氨酸、磷酸苏氨酸）的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

在本发明中，“Glyrichin 蛋白保守性变异多肽”指与表 1 中序列 1~14 的氨基酸序列相比，有至多 10 个，较多地至多 8 个，更佳地至多 5 个，最佳地至多 3 个氨基酸被性质相似或相近的氨基酸所替换而形成多肽。这些保守性变异多肽最好根据表 5 进行氨基酸替换而产生。

表 5 Glyrichin 蛋白可取代氨基酸列表

最初的残基	代表性的取代	优选的取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro; Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe	Leu
Leu (L)	Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala	Leu

本文所用“简并的变异体”在本发明中是指编码具有表 1 中序列 1~14 的蛋白质，但与表 1 中序列 1~14 中相应编码区序列有差别的核酸序列。目前，已经可以完全通过化学合成为得到编码本发明蛋白（或其片段，或其衍生物）的 DNA 序列。然后可将该 DNA 序列引入本领域中已知的各种现有的 DNA 分子（或如载体）和细胞中。此外，还可通过化学合成将突变引入本发

明蛋白序列中。

在本发明中，术语“Glyrichin 家族成员”、“Glyrichin 蛋白”或“Glyrichin 多肽”可相互换使用，都指具有天然抗菌肽 Glyrichin 家族各成员氨基酸序列(表 1 中序列 1-14)的蛋白或多肽。它们包括含有或不含起始甲硫氨酸的天然抗菌肽 Glyrichin 家族成员，以及含有或不含有信号肽的 Glyrichin 成员蛋白。

本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的载体或 Glyrichin 蛋白编码序列经基因工程产生的宿主细胞，以及经重组技术产生本发明所述多肽的方法。通过常规的重组 DNA 技术，可利用本发明的多聚核苷酸序列表达或生产重组的 Glyrichin 多肽。一般来说有以下步骤：

(1) 用本发明的编码 Glyrichin 多肽的多核苷酸(或变异体)，或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞；

(2) 在合适的培养基中培养的宿主细胞；

(3) 从培养基或细胞中分离、纯化蛋白质。

本发明中，Glyrichin 多核苷酸序列可插入到重组表达载体中，术语“重组表达载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病逆转录病毒或其它载体。总之，只要能在宿主体内复制和稳定，任何质粒和载体都可以用。表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。

本领域的技术人员熟知的方法能用于构建含 Glyrichin 编码 DNA 序列和合适的转录/翻译控制信号的表达载体。

包含上述的适当 DNA 序列以及适当启动子或者控制序列的载体，可以用于转化适当的宿主细胞，以使其能够表达蛋白质。

本领域普通技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。获得的转化子可用常规方法培养，表达本发明的基因编码的多肽。在上面的方法中的重组多肽可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要，可利用其物理的、化学的和其它特性通过各种分离方法分离和纯

化重组的蛋白。在使用本发明 Glyrichin 蛋白时，还可同时使用其它药剂，如青霉素等抗菌素。

发明人通过实验，首次发现了 Glyrichin 蛋白家族，得到了其保守序列，并证明了 Glyrichin 蛋白具有能抑制细菌生长的作用。Glyrichin 蛋白可以应用于医药领域，以及食品防腐，鲜花保鲜，化妆品，种子包衣和动物饲料添加剂等方面，使用本发明的 Glyrichin 蛋白家族进行抗菌，具有抗菌活性高，抗菌谱广，靶菌株不易产生抗性突变等特点，有广阔的应用前景。

附图说明

图 1 为 Glyrichin 家族成员同源性比较图，左边为基因来源及注册号，右边为相应序列。

图 2 为 Glyrichin 家族成员同源性比较中的 12 条序列进化树分析图。

图 3 为保守的 Glyrich 结构域，保守的氨基酸用反显字表示。

图 4 为 Glyrich 结构域在不同成员间的序列一致性。

图 5 为 Glyrichin 家族的结构特点图。

图 6 为人源 Glyrichin Northern blot 分析和体外转录与翻译分析，其中

图 6A 为以 hGlyrichin 为探针检测其转录本大小和表达谱，

图 6B 为人源 Glyrichin 体外转录和翻译试验的结果。

图 7 为内源性诱导表达的 h/mGlyrichin 抑制了大肠杆菌 BL21 的生长图。其

中

图 7A 为 pET-22b (+) 空载体，

图 7B 为 pET-22b (+) -UBF 和 pET-22b (+) -PTP，

图 7C 为 pET-22b-hGlyrichin 阳性克隆 1 和 8；

图 7D 为 pET-22b-hGlyrichin 阳性克隆 1 的两次重复实验结果。

图 8 为 hGlyrichin 原核表达产物纯化的 PAGE 电泳结果。其中泳道 1 为蛋白质分子量标准，泳道 2 为空载体诱导表达，泳道 3 为空载体未诱导表达，泳道 4 为 hGlyrichin 克隆基因诱导表达，泳道 5 为 hGlyrichin

克隆基因未诱导表达。

具体实施方式

下面以人源 Glyrichin 基因的克隆、表达、纯化和抗菌活性测试实施例为例，进一步阐述如何利用 Glyrichin 家族各成员本发明。应理解，这些实施例仅用于说明本发明而不用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如 Sambrook 等人的分子克隆：实验室手册 (New York: Cold Spring Harbor Laboratory Press, 1989) 中所述的条件，或按照制造厂商所建议的条件。

实施例 1 人源 Glyrichin 基因的制备及检测

1、人源 Glyrichin 基因完整 ORF 序列的获得

根据鼠源 Glyrichin 基因设计如下一对简并引物扩增人胎肝 cDNA 文库，获得人源 Glyrichin 完整的 ORF 序列共 240bp，以及根据 ORF 序列推测的氨基酸序列：

5' 引物 Pa: 5-CGATGCCGGTGGCCGTGGGTCCCT-3

3' 引物 Pb: 5-TTAGCATCGTATGCCATTCCA-3

PCR 扩增条件为：94°C、4 分钟、1 循环；94°C、40 秒，60°C、50 秒，72°C、1 分钟，30 循环；72°C、7 分钟、1 循环。PCR 产物经过 Winzard PCR preps purification kit 纯化，用 T4 连接酶重组到 pGEM-T 载体中，然后转化大肠杆菌 JM109，测序鉴定。

2、Northern blot 分析

培养人 4 种肿瘤细胞株 HepG2、HeLa、Jurkert 和 HEK293，然后用 Winzard plus RNA purification kit 提取总 RNA。各取 20 微克总 RNA，在 1.2% 甲醛变性琼脂糖凝胶上分离，并转到 Hybond-N+ 尼龙膜。HGlyrichin 探针用 Promega 公司的 Prime-a-gene 试剂盒标记。杂交结果表明 Glyrichin 基因在所测试的 4 种不同组织来源的肿瘤细胞株中都有表达；同时揭示只存在一个转录本，且大小约为 600bp。

3. 体外转录和翻译试验

通过体外转录和翻译试验，证实了 hGlyrichin 基因编码的蛋白体外翻译后的大小约为 8.8kDa，同理论推断的结果一致。将 hGlyrichin 全长的 ORF 序列通过下列引物，PCR 扩增获得。

5' 引物: 5' -CGGGATCCCGATGCCGGTGGCCGTGGGTCCCT-3'

3' 引物: 5' -gctcgagtttagcatcgatgccatcc-3'

然后用 BamHI 和 SacI 酶切；将 pT7 载体用同样的酶切割，然后用 T4 DNA 连接酶连接，构建成功 pT7-hGlyrichin 表达载体。环状的载体在如下体系中进行转录和翻译反应：

25μl 体系：

兔网织红细胞裂解液：12.5μl

反应缓冲液： 1.0μl

氨基酸混合物： 0.5μl

S35-Mat: 1.0μl (50μl i)

DN 2.0μl (0.5μg/μl)

RNase inhibititon: 0.5μl

T7 polyase: 0.5μl

去离子水: 7.0μl

总体积: 25.0μl

30℃ 90 分钟反应。反应结束后取 5μl 样品，加入 10μl Loading Buffer，进行 SDS-PAGE 电泳，固定液中固定 30 分钟，干燥剂中浸泡 5 分钟，用干燥架固定凝胶，干燥凝胶过夜。放射自显影，-20℃ 压片 24 小时，洗片，观察结果如图 6B。

实施例 2 内源性诱导表达的 h/mGlyrichin

对大肠杆菌 BL21 的生长抑制试验

将 hGlyrichin 不带任何标签（即不同任何蛋白融合）重组到原核表达载体 pET-22b (+) 的 EcoR I 和 Xho I 位点之间，转化大肠杆菌 BL21，通过酶切鉴定出阳性克隆。挑取阳性克隆 1 和 8，接种到 AMP 抗性的液体 LB 培养基中，于 37℃ 摆床过夜；次日以 1: 100 接种，继续摇菌至 OD 值为 0.03 时，向每只试管中加入 IPTG 至终浓度为 0.5mM，阴性对照组则加入相应体积的

PBS；于 30℃、250rpm 摆菌，每隔 45 分钟取出 1 毫升菌液测 OD600 值，连续测 20 次。以时间为横轴，OD600 对数值为纵轴做生长曲线。图 7A 为转化了 pET-22b (+) 空载体的 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线；图 7B 为转化了不具抗菌活性的无关基因 pET-22b (+) -UBF 和 pET-22b (+) -PTP 的 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线；图 7C 为转化了待测的 pET-22b (+) -hGlyrichin 阳性克隆 1 和 8 的 BL21 在 0.5mM IPTG 诱导和不诱导条件下的半对数生长曲线；图 7D 为待测的 pET-22b (+) -hGlyrichin 阳性克隆 1 的两次重复。实验结果表明在转入 Glyrichin 后，在 IPTG 诱导表达 Glyrichin 后，细菌的生长明显受到抑制；随着培养时间延长 IPTG 消耗殆尽，细菌的生长又得以恢复，而其他测试基因都没有此现象，充分说明 Glyrichin 蛋白本身具有抗菌活性。

实施例 3 游离 hGlyrichin 表达产物的纯化及杀菌活性检测

根据小鼠 Glyrichin 的 cDNA 序列设计引物，在引物中引入 Enterokinase 的酶切位点，以人胎肝 mRNA 进行 RT-PCR 扩增，纯化后 PCR 产物经测序和限制性内切酶消化后在 T4 连接酶的作用下，插入到用同样内切酶消化后的 pGEX4-4T2 原核表达载体中，经转化 JM109 大肠杆菌后，以 IPTG 进行体外诱导表达 5 小时，收集菌体；用 PBS 重悬，反复冻融裂解；于 4℃、12000rpm 离心 20 分钟，取上清进行 SDS-PAGE 电泳。确定 GST-Glyrichin 融合蛋白在上清中表达后，用 GST 亲和柱纯化，然后用中酶切下来的游离 hGlyrichin 蛋白具有杀菌活性。如上所述诱导表达得到 GST-Glyrichin 融合蛋白，经 GST 纯化柱纯化后用 Enterokinase 切割，得到游离的 hGlyrichin 蛋白，用于最小抑菌浓度 (MIC) 测定。在抗菌活性测定中，稀释菌液的浓度为 10^4 – 10^5 CFU/ml，按每孔 80 微升菌液接种于 96 孔板中，将多肽以一定比例稀释，每孔加入 5 微升；将 96 孔板于 37℃ 培养过夜，紫外分析仪检测 OD600 值，结果如图 8 所示，游离的 Glyrichin 具有在较低浓度下抑制细菌生长的活性，而 GST-Glyrichin 融合蛋白没有抑制细菌生长的活性。

说 明 书 附 图

2/

	10	20	30	40	50
Homo sapiens_CAC11117 79 aa	+	+	+	+	+
0					
Mus musculus_AAK18748 79 aa					
0					
Danio rerio_AAH59661 80aa					
0					
Anopheles gambiae_XP_320383 128aa					YYYYIVVHCCDN
12					
Drosophila melanogas_NP_648757 79aa					
0					
Caenorhabditis elegans_NP_496804 145aa					
0					
Caenorhabditis elegans_B88349 162aa					
0					
Schizosaccharomyces pombe_CAB89006 120a					
0					
Saccharomyces cerevisiae_AAB68201 113aa					
0					
Arabidopsis thaliana_NP_566324 74 aa					
0					
Plasmodium falciparum 3D7_NP_705421 168	MMNWFRKKDSNENKKVKSEYDEYVTPPPFGNYLVSEP				KKPKS-
42					
Plasmodium yoelii yoelii_EAA17587 167 a	MMNWFKKKETTEEPQVKSEYDSYVTPPPFGNYLAKKP				EKPKS-
42					

	60	70	80	90	100	110
Homo sapiens_CAC11117 79 aa	+	+	+	+	+	+
10					MP-VAVGPYQG	
Mus musculus_AAK18748 79 aa					MP-VAVGPYQG	
10						
Danio rerio_AAH59661 80aa				MP-VSVGSYCQQ		
11						
Anopheles gambiae_XP_320383 128aa		THFNEFVPKIKLPPRKRYVRSGSFQILQKTDKSTMPAVPGGVYSQN				
59						
Drosophila melanogas_NP_648757 79aa				MP-LPTSSPSQ		
10						
Caenorhabditis elegans_NP_496804 145aa				MP-VPSGYAAH		
10						

22
Caenorhabditis elegans_B88349 162aa -----MQ-HTHKEANRRVLQRKKIN
19

Schizosaccharomyces pombe_CAB89006 120a -----MQSM-----
4

Saccharomyces cerevisiae_AAB68201 113aa -----MPPLPQNYAQ-----
11

Arabidopsis thaliana_NP_566324 74 aa -----MAKN-----
4

Plasmodium falciparum 3D7_NP_705421 168 -----LKNDKTAITEFKGFTPPPKFEFKEDISDNKYEEDFSKYTSNNIIDSSFYDD
93

Plasmodium yoelii yoelii_EAA17587 167 a -----LKNEKINVTEFKGFTPPPKFEFKEDTTDTQYDQDFSKYTNNNFIDSSFYDD
93

120 130 140 150 160

Homo sapiens_CAC11117 79 aa -----SQPSC-----FD--RVKMGPVMGCAVGMAAGALPGTFSCLRIGMRGREL MG
54

Mus musculus_AAK18748 79 aa -----SQPSC-----FD--RVKMGPVMGCAVGMAAGALPGTFSCLRIGMRGREL MG
54

Danio rerio_AAH59661 80aa -----AQPSC-----FD--RVKMGFMMGFAVGMAAGAMFGTFSCLRIGMRGREL MG
55

Anopheles gambiae_XP_320383 128aa -----QQPSC-----FD--RMKMGPTIGFCVGMASGALFGGFSALRYGLRGRELIN
103

Drosophila melanogas_NP_648757 79aa -----QGPTC-----FD--KMKTGFIIGFCVGMASGAVFGGFSALRYGLRGRELIN
54

Caenorhabditis elegans_NP_496804 145aa -----PQGSQPSC-----FT--KIRMGLMMGAMIGGATGILLGGFMGFRAGMRGKDLLL
57

Caenorhabditis elegans_B88349 162aa -----LLEMSDKICRNLIYFQNFQIRMGLMMGAMIGGATGILLGGFMGFRAGMRGKDLLL
74

Schizosaccharomyces pombe_CAB89006 120a -----QPST-----VD--KLKMGAIMGSAAGLIGFLFGGVAVLRYGPGPRGFLR
47

Saccharomyces cerevisiae_AAB68201 113aa -----QPSN-----WD--KFKMGLMMGTTVGVCTGILFGGFAITQGPGPDGVVR
54

Arabidopsis thaliana_NP_566324 74 aa -----SC-----LA--KITAGVAVGGALGGAVGAVYGTYEAIRVKVPGHLKVR
45

Plasmodium falciparum 3D7_NP_705421 168 KKLLSDVNLSHRTRACFESIKMGVKMGTMVGGIFGSLTGIYASFAHKN-----LF
143

Plasmodium yoelii yoelii_EAA17587 167 a KPNMFDFTLSHRTKACLESVKMGVKMGTMVGGIFGSLTGLYASFAHKN-----LF
143

170 180 190 200 210 220

Homo sapiens_CAC11117 79 aa -----GIGKTMMQSGGTFGTFMAIGMGIRC
79

✓ 3

Mus musculus_AAK18748 79 aa	GIGKTMMQSGGTFGTFMAIGMGIRC
79	
Danio rerio_AAH59661 80aa	GVGKTMMQSGGTFGTFMAIGMGIRC
80	
Anopheles gambiae_XP_320383 128aa	NVGKVMVQGGTFCGTMAIGTGIRC
128	
Drosophila melanogas_NP_648757 79aa	NVGKTMVQGGTFCGTMAIGTGIRC
79	
Caenorhabditis elegans_NP_496804 145aa	QTGKTVAQSGGSFGVFMGVAQGLRYIFFKNLAGTGFWPFSLNFSRSIDNCPRGV
112	
Caenorhabditis elegans_B88349 162aa	QTGKTVAQSGGSFGVFMGVAQGLRYIFFKNLAGTGFWPFSLNFSRSIDNCPRGV
129	
Schizosaccharomyces pombe_CAB89006 120a	TLGQYMLTSAAATFGFFMSIGSVIRNEDIP-LIQQSGSHWNQRLLNENAN-----
95	
Saccharomyces cerevisiae_AAB68201 113aa	TLGKYIAGSAGTFGLFMSIGSIIRSDSE---SSPMSPHPNLNLQQAR-----
98	
Arabidopsis thaliana_NP_566324 74 aa	FIGQTLSSAAIFGLFLGAGSLIHCG-----KGY
74	
Plasmodium falciparum 3D7_NP_705421 168	ILPVSVLGGAVSPGFFLGCGMIVRC
168	
Plasmodium yoelii yoelii_EAA17587 167 a	IFPVSVIGGAVSFGFFLGCGMIVR
167	

230	240	250
-----	-----	-----

Homo sapiens_CAC11117 79 aa	
79	
Mus musculus_AAK18748 79 aa	
79	
Danio rerio_AAH59661 80aa	
80	
Anopheles gambiae_XP_320383 128aa	
128	
Drosophila melanogas_NP_648757 79aa	
79	
Caenorhabditis elegans_NP_496804 145aa	VYSTRNAFRFTTEIEKKEPRRDTQAVNLPQI
145	
Caenorhabditis elegans_B88349 162aa	VYSTRNAFRFTTEIEKKEPRRDTQAVNLPQI
162	
Schizosaccharomyces pombe_CAB89006 120a	--SSRI--FALAMQQAKSSPRKSNEVAEC
120	
Saccharomyces cerevisiae_AAB68201 113aa	-----LEMWKLRAKYGIRKD
113	
Arabidopsis thaliana_NP_566324 74 aa	
74	

Plasmodium falciparum 3D7_NP_705421 168

168

Plasmodium yoelii yoelii_EAA17587 167 a

167

24

图 1

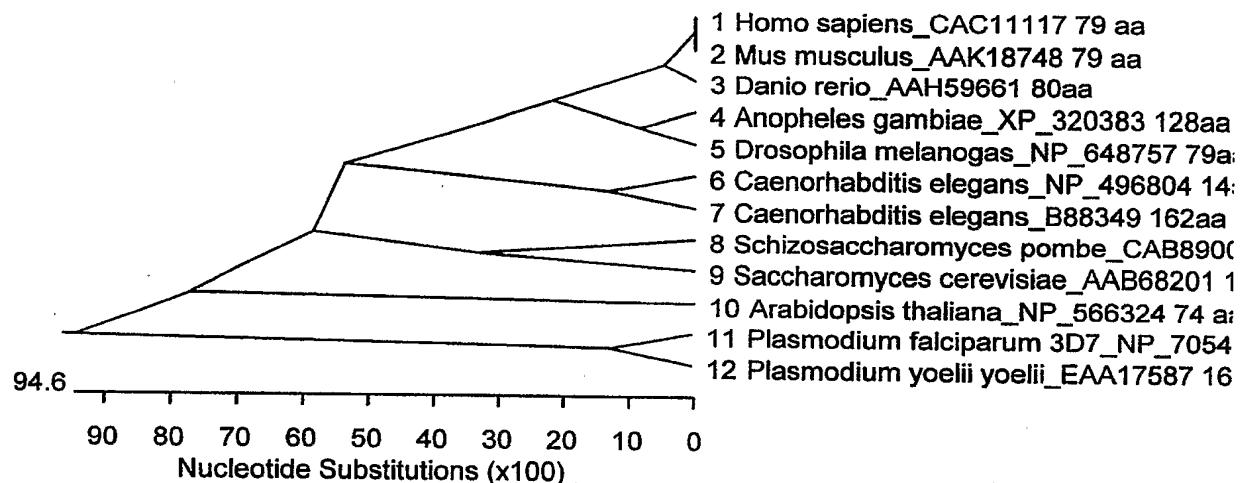


图 2

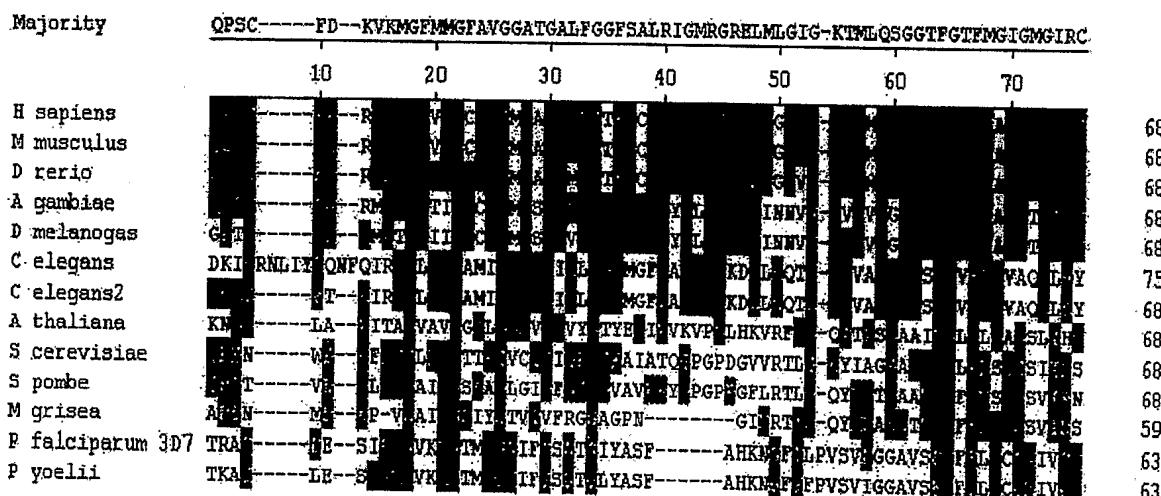


图 3

25

		Percent Identity														
		1	2	3	4	5	6	7	8	9	10	11	12	13		
Divergence	1	100.0	94.7	76.9	74.1	36.3	42.1	50.0	51.3	40.8	35.6	35.5	42.1	4		
	2	100.0	94.7	76.9	74.1	35.3	42.1	50.0	51.3	40.8	35.6	35.5	42.1	2		
	3	67.1	61.1	74.6	73.7	35.3	42.1	50.0	50.0	39.3	32.9	32.9	42.1	3		
	4	32.7	32.7	30.4	30.6	31.3	36.2	38.7	50.0	35.6	32.9	31.6	38.2	4		
	5	42.2	42.2	37.3	11.1	19.7	36.5	44.7	47.4	38.2	30.3	28.9	39.5	5		
	6	79.9	19.9	78.0	92.5	102.0	84.2	48.7	40.6	36.8	38.2	35.5	36.8	6		
	7	92.5	92.5	92.5	107.1	112.6	7.8	34.2	26.3	25.0	28.9	27.6	23.7	7		
	8	97.1	97.1	102.0	118.1	112.0	44.6	63.2	35.6	35.5	35.5	44.7	8			
	9	92.5	92.5	97.1	99.1	104.1	104.4	74.0	50.0	38.6	32.9	31.6	51.3	9		
	10	138.4	130.4	146.1	174.0	154.5	163.9	185.4	174.0	146.1	39.5	38.2	34.2	10		
	11	146.8	146.6	166.3	166.3	191.1	1130.3	130.3	146.6	166.3	173.4	92.1	20.3	11		
	12	146.8	148.6	166.3	177.9	202.0	146.6	130.3	148.6	177.9	130.3	10.7	20.9	12		
	13	108.9	108.9	108.9	129.8	122.3	128.0	167.8	97.3	73.3	156.8	194.2	207.0	13		
		1	2	3	4	5	6	7	8	9	10	11	12	13		

图 4

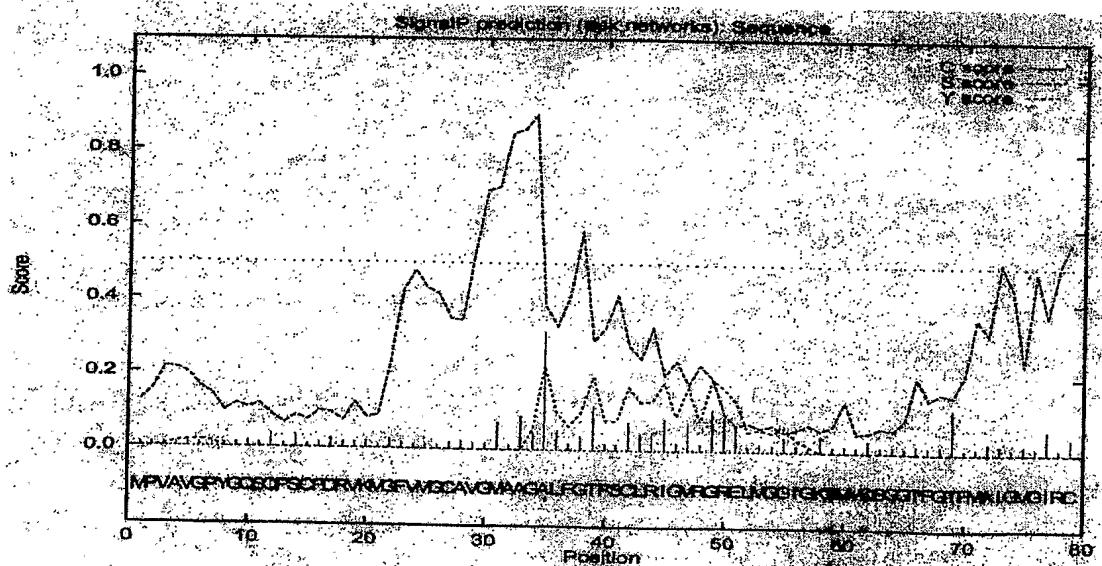
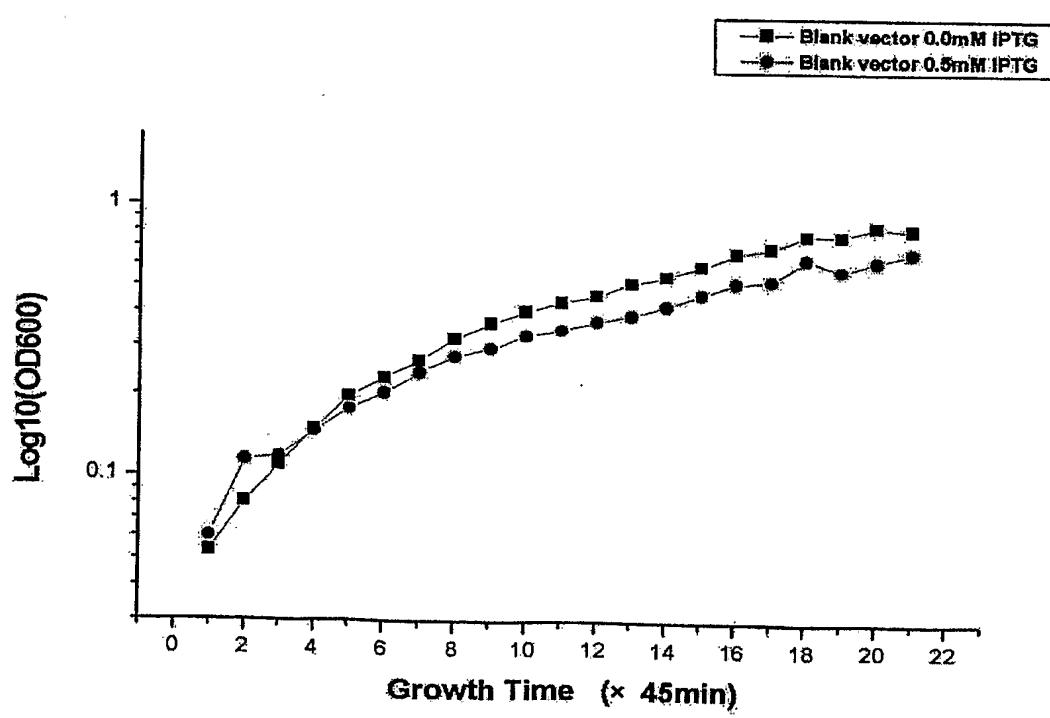
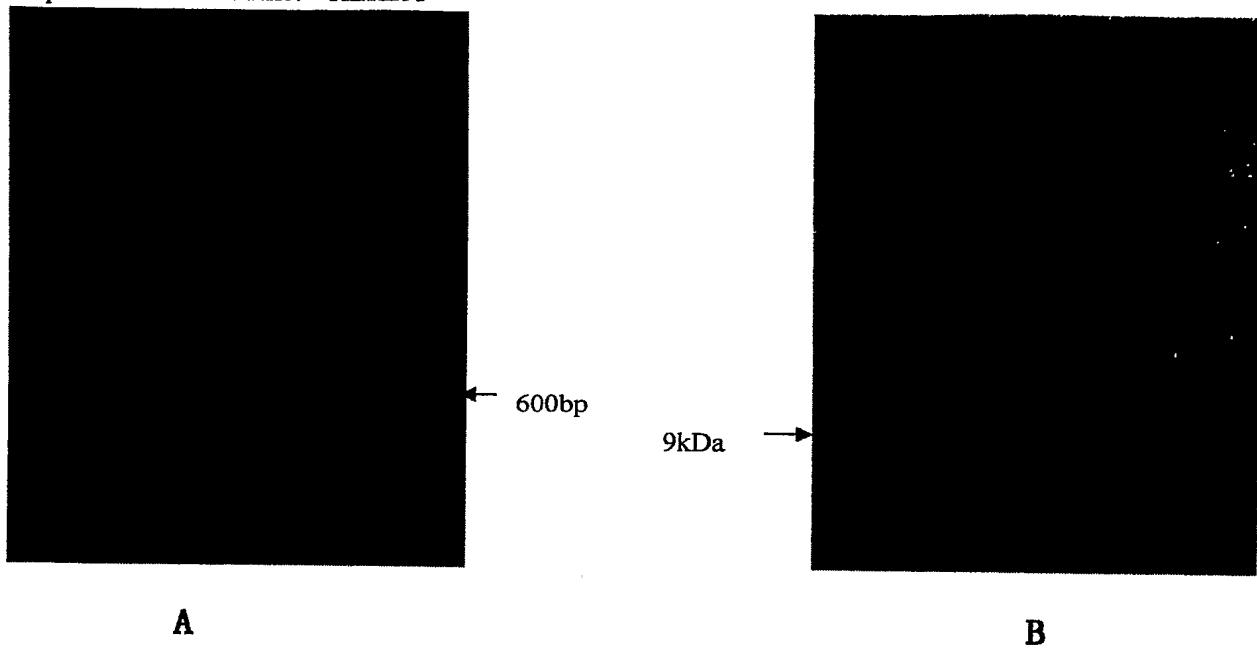
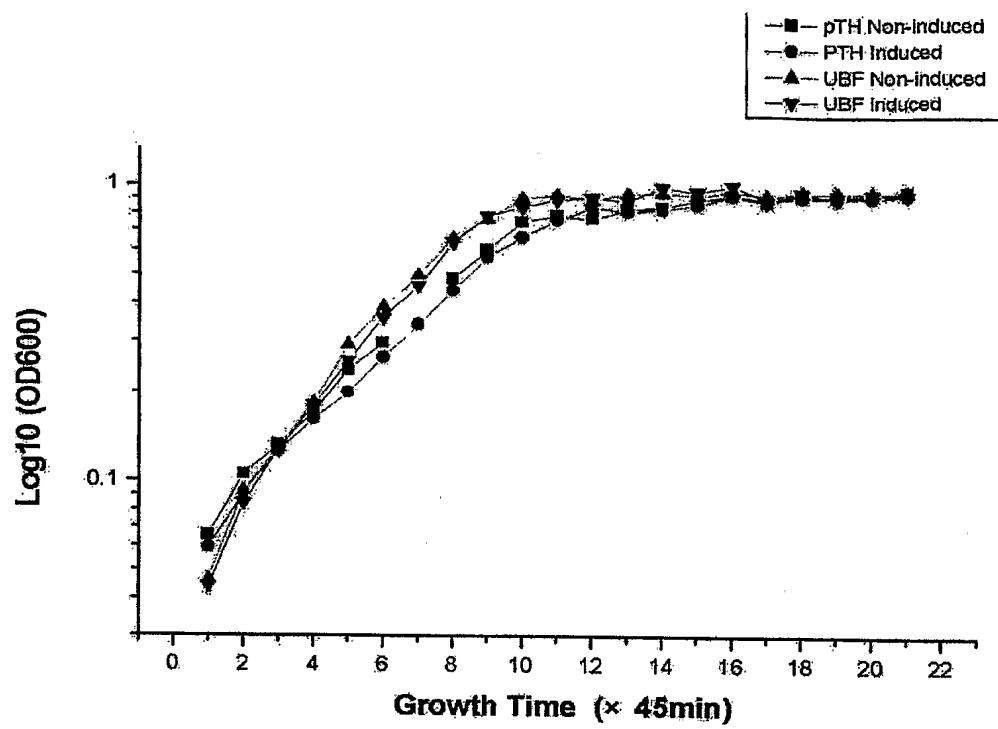


图 5

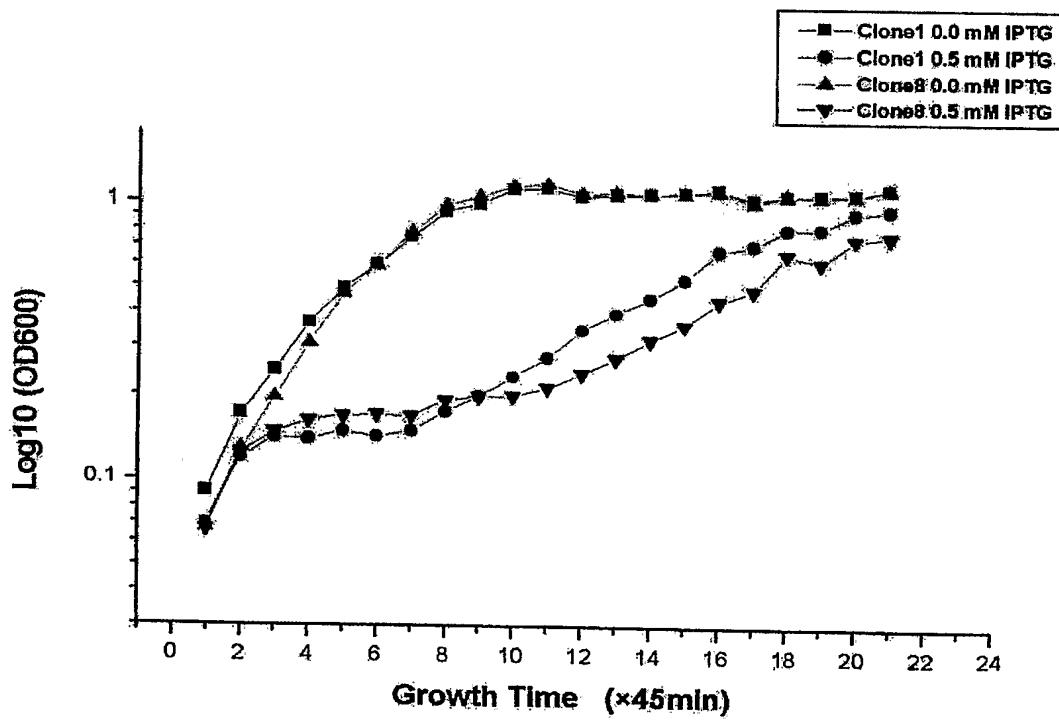
26
HepG2 HeLa Jurket HEK293



A



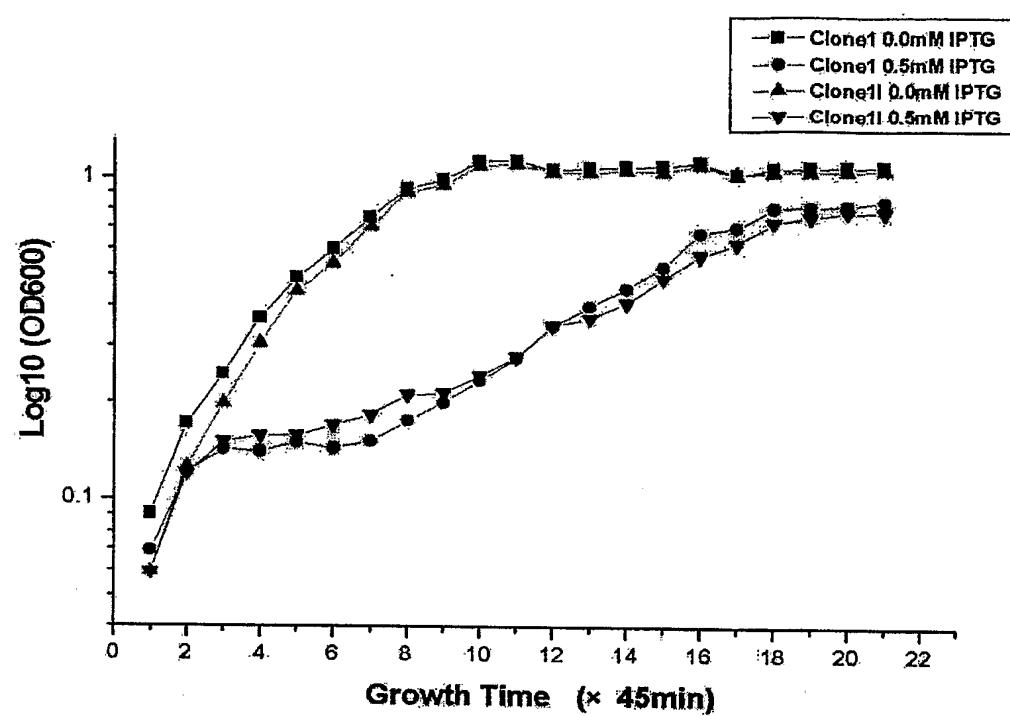
B



C

000-112-11

28



D

图 7

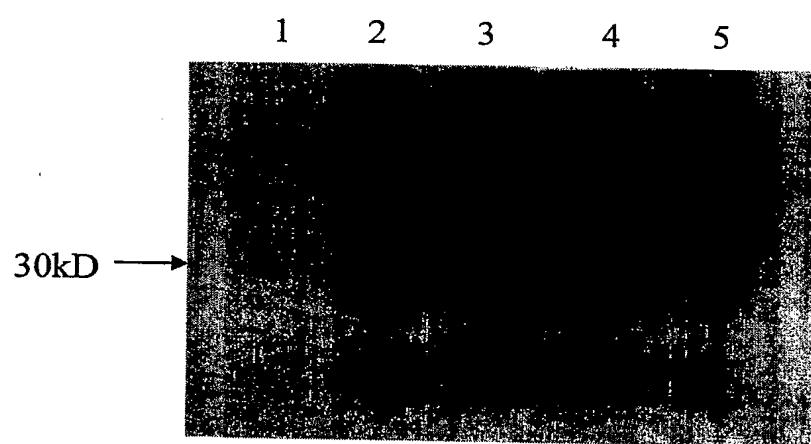


图 8